

המרכז
האקדמי
לב

מודיעין עסקי תחרותי

רכז הקורס: אורי שטרנברג

יחידה 4 - חיפוש מידע ברשת האינטרנט

* מי מעלה מידע לאינטרנט?

* אנשים פרטיים



* חברות ציבוריות ופרטיות



* ארגונים שונים



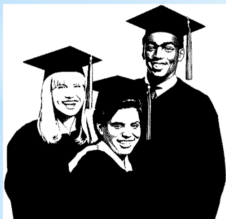
* ממשלה



* צבאות



* אוניברסיטאות ומכוני מחקר



ועוד...

*עיוורון רשת

* קיימת יותר נגישות למידע, אך קשה יותר למצוא מידע **רלוונטי**, מכיוון שהמידע מפוזר ללא הקשר, התוכן לקטני (אקלקטי) ואיכות המידע ירודה.
* התיישנות מהירה של מידע, מה שמוביל לכך שקיים מידע רב שאינו רלוונטי יותר, אך מצוי בשטח - לעיתים בזמינות גבוהה ממידע חדש יותר.
* האתגר:
לסנן את המידע (filter feeding), שכן באינטרנט קיימים כ-50 מיליארד דפים במעל מאה מיליון אתרי אינטרנט.
האתרים ודפי האינטרנט הינם בפורמטים שונים.
כל 8 חודשים לערך האינטרנט מכפיל את כמות המידע שבו. בכל יום נוספים כ 2.5-3 מיליון דפים חדשים.



ארגון המידע ברשת האינטרנט

- רשת האינטרנט (ה-Web) אינה מנוהלת כקטלוג ספריה. אין אחידות בדרכי ארגון ופרסום המידע.
  - המידע באינטרנט מאורגן במבנה של אתר - Site
  - אתר מורכב ממסמכים - דפים - Pages
  - קישור בין דפים - link
  - מעבר בין דף לדף נקראת גלישה - Surfing
  - אחזור מידע נעשה עם היפר-טקסט - Hypertext

*מבנה כתובת באינטרנט (URL)

* המבנה מורכב בדרך כלל מ-4 אלמנטים:
  * ציון סוג האתר (פרוטוקול ההתקשרות והרשת)
  * שמו (הדומיין שלו)
  * סוג הבעלות על האתר (ממשל, אקדמי, עסקי...)
  * ארץ המוצא או בינלאומי

http://www.tnuva.co.il/

HTTP– פרוטוקול התקשרות

WWW - מציין שזהו דף WEB

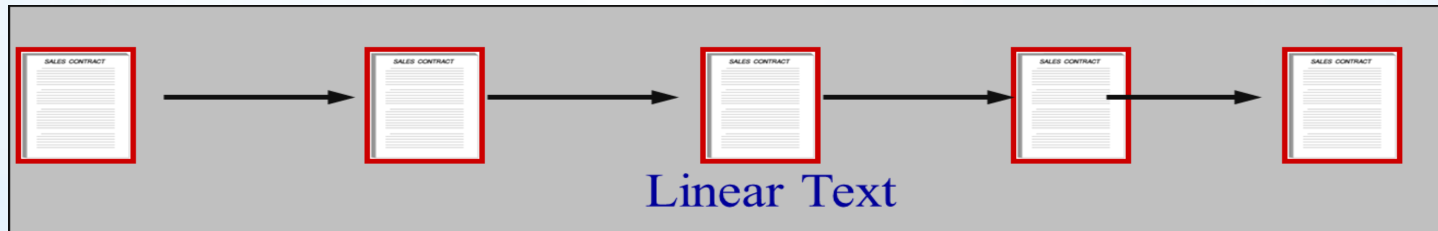
Domain - **tnuva**

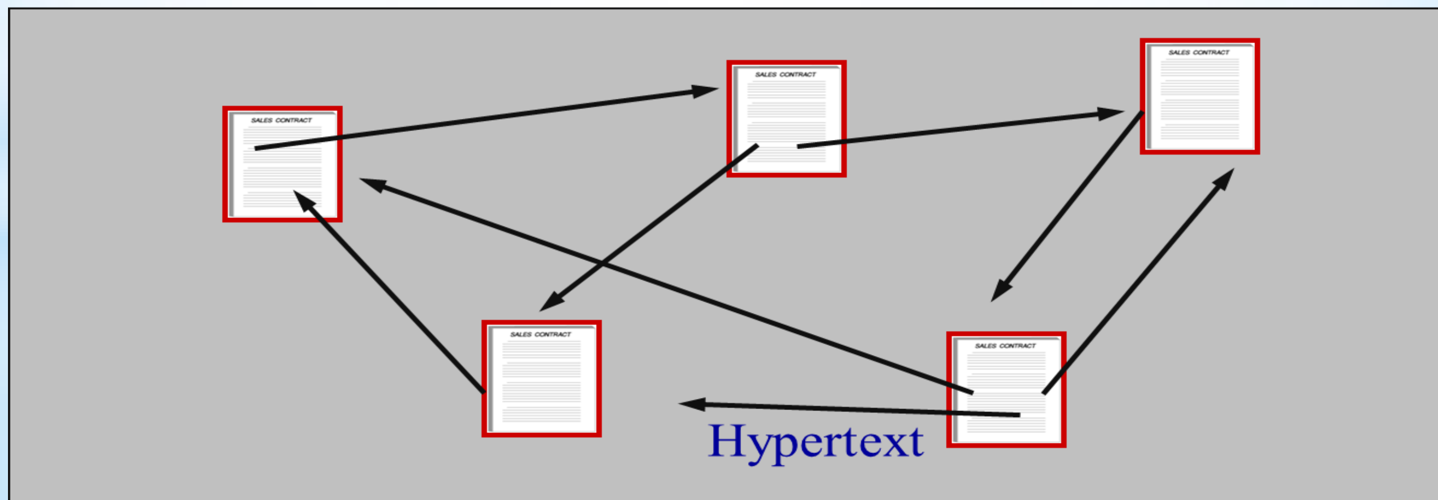
co - מציין שזו חברה

il - ישראל

*שימוש בהיפר טקסט (Hypertext)

* ניתן לכתוב דפי אינטרנט בשתי דרכים:
  * ליניארי - ללא שימוש בהיפר-טקסט
  * מתקדם תוך שימוש בהיפר-טקסט. הצגת כמויות טקסט בדרך מקבילה.





*דוגמה גרפית לשימוש בהיפר טקסט (Hypertext)



המארג הכלל עולמי (מרשתת) World Wide Web

* המארג מכיל כמיליארד צמתים המקשרים בין מסמכים (דפי אינטרנט).
* במחקר שערכו סטיב לורנס ולי ג'יילס ממכון המחקר NEC אשר בפרינסטון, נמצא כי כל מסמך (דף אינטרנט) נמצא במרחק ממוצע של 19 לחיצות מכל מסמך (דף אינטרנט) אחר.
  * עולה השאלה: איך הרשת מצליחה לייצר מסלול כל כך קצר בין מסמך למסמך?
    * התשובה נובעת מרמת הקישוריות הגבוהה של הרשתות.
      * מספרם הממוצע של הקישורים בכל מסמך ב- WWWהוא בערך שבעה.
      * הדבר אומר שבעוד אנחנו יכולים לעקוב אחר שבעה קישורים בלבד מהמסמך הראשון, ישנם 49 קישורים במרחק שתי לחיצות, 343 במרחק שלוש לחיצות וכך הלאה.
      * כך, נגיע לצמתים הנמצאים במרחק 19 קישורים רק לאחר שנבדוק 10 מסמכים בחזקת 16- פי עשרה מיליון ממספר הדפים במארג...

* המארג הכלל עולמי (מרשתת) World Wide Web

* שאלה נוספת שעולה עוסקת ב**אקראיות** של המארג.
  * ההשקפה ששלטה משך עשרות שנים, אשר לפיה הרשתות הן אקראיות, נבחנת בימים אלו מחדש...
  * המודל החדש מציע הסבר פשוט של התקבצות לצבירים, המכנס את הרשתות האקראיות למקבצים בעלי מכנה משותף כלשהו.
  * כך, מיפוי הצבירים גורם למארג להיראות כמו 'עולם קטן' בעל היגיון הדומה לעולמנו שלנו.
  * אם קיימים קישורים למספר גדול של צמתים, בצבירים נוצרים מסלולים קצרים בין כל שני צמתים במערכת.
  * כתוצאה מכך, בעוד שכל שני מסמכים במארג נמצאים במרחק 19 לחיצות זה מזה, אפשר להגיע באמצעות צמתים למרבית המסמכים ברשת באמצעות 2-3 לחיצות בלבד.



דוגמה גרפית לצביר מידע המכיל מספר צמתים

Mozilla Firefox
Debian GNU/Linux
Bobby
Microsoft Internet Explorer
WordPress
mozilla.org
PHP-Nuke
Viewable With Any Browser
W3C CSS Validator
Macromedia France - Centre de
The Online Books Page
Opera Software
AddAll
Sun Microsystems
Compaq Product Information
WinZip
IAF.net - Internet Address Finder
Netscape.com
Microsoft Corporation
White Pages listings on WhoWhere?
Macromedia
Adobe Systems Incorporated
Welcome to Bibliofind
BookFinder.com
IBM Corporation
Apple Computer, Inc.
Apple Quicktime
abebooks.com
Switchboard
Librarians' Index to the Internet - lii.org
uBid
Washington Post
Borders teamed with Amazon.
Search Engine Watch
Vivisimo Clustering Engine
Mamma
MSN Search
Barnes and Noble
Startpagina.nl
Ixquick Metasearch
Open Directory Project
The Weather Channel
Amazon.com: Welcome
Google Groups
AllTheWeb
Teoma
Search.com
Google
Ask Jeeves
CDNOW
My Excite
Yahoo!
GO.com
My Way
CNN - Cable News Network
helpdesk
AltaVista
USA Today
PayPal
The Internet Movie Database
Powell's Books
LookSmart
http://search.aol.com/
Dictionary.com
Movable Type
The New York Times
KidsClick! Web Search
PBS Kids
Boing Boing
FindLaw
Fark.com
The Register
Encyclopaedia Britannica
Babelfish
OneLook Dictionary Search
Blogger.com
Freshmeat
User Friendly
Blogger: 404 - Page not found
KuroShin.org
Blogwise
HaloScan
Red Hat, Inc.
Blogarama
Tag-Board Smilies
Hello : Welcome

*עבודה עם צבירים וצמתי מידע

* לא עוקבים אחר כל הקישורים, אלא מחפשים רמזים.
* מספקים פרשנות לקישורים.
  * האם מדובר בקשר הדוק או בקשר חלש ומקרי.
  * איתור צמתי מידע - אתרים עם הרבה קישורים.



דוגמאות לצמתי מידע באינטרנט

* פורטלים
* מדריכים אינטרנטיים
* אינדקסים
* מנועי חיפוש
* כלי יעץ - מילונים ואנציקלופדיות
* קבוצות דיון
* רשתות חברתיות
* בלוגים
* אתרי מסחר ומאגרי מידע מסחריים

*פורטל (web portal, פורטן)

* שער, כניסה או פתח.
* אתר המהווה נקודת התחלה לגלישה.
* פורטל יכול להוות שער לגלישה בתחום מסוים.
לדוגמה: פורטל טכנולוגי הוא שער לאתרים בנושאים טכנולוגיים.
* פורטלים מעניקים שירותים: תיבת דואר, שירותי מידע בנושאי בורסה, קניות מקוונות, חדשות ועוד.



*דוגמה לפורטל קווי תחבורה

מרכז מידע ארצי
*8787
052-9998787
072-2588787

פורטל קווי תחבורה בישראל
Ministry of Transport - משרד התחבורה



דף הבית | נקה | קווים בין ישובים | צור קשר | עזרה | תנאי שימוש | חדש בקווי התח"צ

26.08.2015 אגד: ירושלים, קו 136 לקבר שמואל הנביא - שינוי סטטוס תחנות
26.08.2015 אגד: ירושלים, קו 16 - תוספת תחנה בשדרות הרצל

מוצא ויעד | מס' קו | שפה חופשית

תאריך 29/08/2015 - יום ש משעה הכל
מוצא, תחנת עליה
ישוב
מקום בישוב מיון לפי חשיבות א'ב'
חיפוש לפי כתובת
יעד, תחנת ירידה
ישוב
מקום בישוב מיון לפי חשיבות א'ב'
חיפוש לפי כתובת
חפש | הפוך כיוון

פורטל קווי תחבורה בישראל מספק מידע לנוסע בתחבורה הציבורית,
האתר מציע קווים ולוחות זמנים של כל חברות האוטובוסים ורכבת ישראל
כולל שילוב קווים של מפעילים שונים.
תכנן את נסיעתך באתר אחד.
תוכל לחפש לפי מוצא ויעד, לפי מס קו או בשפה חופשית.
באתר מידע של החברות: אגד,אגד תעבורה,אפיקים,ג.ב. טורס,גלים,דן,
כרמלית,מטרודן,מטרופולין,נסיעות ותיירות,נתיב אקספרס,סופרבוס,
עילית, קווים, ויאוליה ,רכבת ישראל, רכבת קלה בירושלים, ש.א.מ

חניוני חנה וסע חדשים, פרטים באתר: www.busline.co.il

אם לא יצוינו תחנות מוצא ויעד מדויקות, המידע יתייחס למקום מרכזי בישובי
המוצא והיעד. לשם קבלת מידע נוסף על קווים ותעריפים בין ישובים
ניתן להסתייע ב קווים בין ישובים.

בפעילות נוספת באתר יסגר חלון זה.

סגור ✕ עזרה

* דוגמה לפורטל דרושים

פורטל דרושים

Like 389 · 6,024 g+1

אין משרות בסל | שלום אורח, התחבר | הרשם

דרושים | דרושים הייטק | כל הקטגוריות | קורות חיים | סוכן חכם | צור קשר | איזור אישי | מעסיקים

כל הקטגוריות ▾ | כל האיזורים ▾ | חפש משרות

אדמיניסטרציה	כללי	משאבי אנוש	רכב / תחבורה	**דרושים הייטק**
ביטוח	כספים	עבודה בחו"ל	רפואה / בריאות	חומרה
בכירים / ניהול	לוגיסטיקה / שילוח	עבודה מהבית	שיווק	כללי
בעלי מקצוע	מדעים / ביוטק	עיצוב	שירות לקוחות	תוכנה
הדרכה / הוראה	מכירות	עריכת דין	שמירה / אבטחה	QA
הנדסה	מסעדנות / תיירות	פרסום / מדיה	תעשיה / ייצור	אינטרנט

מעסיק, מחפש עובדים?

עשרות אלפי מועמדים ביום!

פרסם משרות עכשיו!

חברות מובילות באתר

TOYGA

סופר-פארם

VICTORIA'S SECRET
BEAUTY & ACCESSORIES

הסוכן החכם
לקבלת משרות באימייל הרשמו כאן
הזן כתובת אימייל
בחירת קטגוריה ▾
הצטרף חינם

טבלאות שכר
נתוני שכר מעודכנים לפי תחום

קורות חיים
קבלו קורות חיים מקצועיים, ובחינם!

מחשבון שכר
מברוטו לנטו

מידע ומאמרים
מדריכים וטיפים לחיפוש עבודה

פורטל דרושים - זה עובד!

מדריך אינטרנטי

* קטלוג של אתרי הרשת, המחולק לקטגוריות ותת קטגוריות המסודרות בצורה היררכית.
* המדריך הוא אינדקס אתרים הנעשה לרוב על ידי בני אדם (ולא תוכנה).
* מכיל כמות קטנה יחסית של אתרים.

מדריך אתרי אינטרנט

הוסף אתר | אתרים חדשים

פנאי ובידור
קולנוע, מוסיקה, מפורסמים, »

ספורט
כדורגל, כדורסל, הרפתקני, »

כלכלה ועסקים
חברות, נדל"ן, שוק-ההון, »

מחשבים ואינטרנט
משחקים, תוכנות, אינטרנט, »

תיירות ונופש
בישראל, בעולם, בתי מלון, »

קניות
מוסיקה, ספרים, מחשבים, »

חדשות ומדריכים
מדריכים, אנציקלופדיות, »

אמנות
ציור, מוזיאונים, אלטרנטיבי, »

חברה וקהילה
הכרויות, משפחה, נוער, »

חוק וממשל
משרדי ממשלה, משפט, »

חינוך ומדע
אוניברסיטאות, בתי-ספר, »

בריאות ורפואה
בתי-חולים וקופ"ח, משלימה, »

פורטל יכול להיות גם מדריך

* בשנת 1994 שני דוקטורנטים להנדסה אלקטרונית פרשו מהלימודים באוניברסיטת סטנפורד והחלו לפתח פרויקט תיעוד חדשני באינטרנט.
* שם הפרויקט המלא: Yet Another Hierarchical Officious Oracle
* כולל מעל 100 מיליון אתרים, המסווגים ומאורגנים בכ - 25,000 קטגוריות נושאיות.

Finance Music Shopping YAHOO! Mail My Yahoo! Messenger

Select Category: Web | Images | Video | Audio | Directory | Local | News | Shopping

Search the Web: Yahoo! Search
Advanced
My Web
Answers

Yahoo! Shopping – Find Valentine's Day Gifts: Jewelry, Gadgets, Flowers, Lingerie, Watches

Check your mail status: Sign In Free mail: Sign Up

360°
Autos
Finance
Games
GeoCities
Groups
Health
Horoscopes
HotJobs
Kids
Local
Mail
Maps
Mobile
Movies
Music
My Yahoo!
News
People Search
Personals
Photos
Real Estate
Shopping
Sports
Travel
TV
Yellow Pages
All Y! Services...

The Coca-Cola Company
make every drop count
Take the Hydration Quiz
© The Coca-Cola Company, All Rights Reserved
Learn about hydration - Open Ad - Ad Feedback

In the News 12:09pm, Thu Feb 9

איך מאתרים מידע?

- איתור צמתי מידע רלוונטיים לתחום החברה.
- איתור פורטלים בתחום הידע של החברה.
- איתור אתרי אינטרנט בנושא החברה או אתרי חברות מתחרות.
- היכרות עם הקטגוריות הרלוונטיות לתחום הידע במדריכי האתרים המסווגים ומאורגנים של קטגוריות היררכיות מקטגוריות על ועד לתת קטגוריות.